

# UC Berkeley

## Proposals from the Script Encoding Initiative

### Title

Final revised proposal to encode the Saurashtra script in the UCS

### Permalink

<https://escholarship.org/uc/item/59q3p5j6>

### Authors

Everson, Michael  
Krishnamoorthy, Jeyakumar Chinnakkonda

### Publication Date

2005-09-21

Peer reviewed

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization  
Organisation internationale de normalisation  
Международная организация по стандартизации

**Doc Type:** Working Group Document  
**Title:** Final revised proposal to encode the Saurashtra script in the UCS  
**Source:** Michael Everson and Jeyakumar Chinnakkonda Krishnamoorthy  
**Status:** Individual Contribution  
**Date:** 2005-09-21

## Introduction

Saurashtra is an Indo-European language, related to Gujarati and spoken by about 310,000 people (1997 India Missions Association) in southern India, mostly in the area around Madurai, Salem, and Thanjavur cities. The SIL *Ethnologue* describes it as “an Indo-European island surrounded by Dravidian languages”. Since the end of the 19th century the Telugu, Tamil, Devanagari, and Saurashtra scripts have been used to publish books in Saurashtra. At present, Saurashtra is most often written in the Tamil script, augmented with the use of superscript digits and a colon to indicate sounds not available in the Tamil script. The Saurashtra script is of the Brahmic type; early Saurashtra text makes use of conjuncts, which can be handled with the usual Brahmic shaping rules. The modern version, developed in the 1880s, has undergone some simplification. Like Tamil, modern Saurashtra does not use complex consonant clusters, marking instead the killed vowel with a visible subscript virama; ્ka is formed of KA + VIRAMA. An exception to this is the conjunct ્kṣa, which is formed of KA + VIRAMA + ZWJ + SSA in this encoding, and which is sorted as a unique letter in older dictionaries. Because, apart from ્kṣa, the virama is always visible in modern Saurashtra, it is best to encode modern Saurashtra with the virama visible by default, and to use ZERO WIDTH JOINER to force conjunct behaviour. Thus this encoding supports both older and newer Saurashtra text. A keyboard specification may make it easy to type ્kṣa with a single keystroke comprising all four characters KA + VIRAMA + ZWJ + SSA.

## Glyph placement

The vowel signs in Saurashtra follow the consonant to which they are applied. The vowels *-i* and *-ī*, however, are typographically joined to the top right corner of their consonant. Examples: ્ka + ી *-i* = ્ki (not ્k̃), ્ka + ુ *-ī* = ્kī (not ્k̄), and ્ṭa + ી *-i* = ્ṭi (not ્ṭ̃), ્ṭa + ુ *-ī* = ્ṭī (not ્ṭ̄). Vowel matras are also applied to the UPAKSHARA (see below).

## Digits and punctuation

Digits have distinctive forms. In the samples viewed at present, European COMMA, FULL STOP, and QUESTION MARK are used, most commonly. Both SAURASHTRA DANDA and SAURASHTRA DOUBLE DANDA are used in poetry; SAURASHTRA DANDA is also used in prose.

## Glyph variants

The shape of the vowel signs for *-l̥* and *-l̄* is noteworthy. In one teaching book it appears that full-size glyphs are used, with a tilde-like mark beneath indicating the long vowel: ્kl̄, ્kl̄̃ (compare ્kr and ્k̄̃). In earlier sources, however, the glyph for *-l̥* is reduced, and length is

marked with  $-\bar{a}$ ; the *Sangudhari* font supports this practice: **ᱦᱚ**  $k\bar{l}$ , **ᱦᱚᱛ**  $k\bar{l}$  (compare *Sangudhari* **ᱦᱚ**  $k\bar{r}$  and **ᱦᱚᱛ**  $k\bar{r}$ ).

# Saurashtra UPAKSHARA

A character ્ called UPAKSHARA, which we transliterate *H* here, is unique to Saurashtra, and does not have an equivalent in the Devanagari, Tamil, or Telugu scripts. It functions in some regards like the Tamil *āyтам*, in that it is used to modify other letters to represent sounds not found in the basic Brahmic alphabet. It differs from *āyтам* because it is a dependent consonant—a kind of consonant matra. In Figure 7, a chart from *Bhāṣābhīmāni*, four Saurashtra letters are shown in combination with the UPAKSHARA: ્*rh*, ્*lh*, ્*mh*, ્*nh*. In this chart, the visible VIRAMA is drawn below the liquids and nasals and before the UPAKSHARA, though in the encoding, it is, for example, RA + UPAKSHARA + VIRAMA. (In that same chart, the Tamil letters are shown with the Tamil *pulli* on two letters: ூ*rh*, ௃*lh*, ௄*mh*, ௅*nh*., a different transliteration and encoding model; here we have RA + VIRAMA + HA + VIRAMA.) Since in running text the visible virama is not shown, it is clear that the UPAKSHARA modifies the consonant as a whole. The related language Marathi also has aspirated liquids and nasal such as these, which are shown as conjuncts with *ha*: ೠ*rha*, ೡ*lha*, ೣ*mha*, ೮*nha*—a different model again. Saurashtra UPAKSHARA functions as an dependent consonant sign, also taking the dependent vowel when used. Here is the vowel repertoire in Saurashtra, Latin, and Tamil:

ညယ	ညၤ	ညၤၤ	ညၢ်	ညၢါ	ညၢၢ	ညၢၢ	ညၢၣ	ညၢၤ	ညၢၥ	ညၢၤ	ညၢၤ	ညၢၤ
<i>mH</i>	<i>mHa</i>	<i>mHā</i>	<i>mHi</i>	<i>mHī</i>	<i>mHu</i>	<i>mHū</i>	<i>mHě</i>	<i>mHe</i>	<i>mHai</i>	<i>mHǎ</i>	<i>mHo</i>	<i>mHau</i>
မံ:	မ:	မာ:	မိ:	မီ:	မု:	မူ:	မဲ:	မေ:	မေး:	မော:	မော:	မော့:

In the publication *Padārthun*, the words **பேழன்** *nHāna* and **பேயன்** *nHīna* are transliterated into Tamil-based orthography as **நா:ந** *nā:na* and **நீ:ந** *nī:na*. **ஐயல** *mHor* and **பேய்ள** *nHannō* are transliterated as **மோ:ர** *mo:r* and **ந:ந்நொ** *na:nnō*. The colon-like mark is used also in Tamil transliterations of this letter in *Bhāṣābhīmāni*, where **பேரிடு** *nHittak* is written **நீ:த்தக்** *nī:ttak*. This colon-like mark is not encoded in the Tamil block; one could consider it a new sort of VISARGA, perhaps. Another example of a word written with the UPAKSHARA is **பேய்ள** *nHīnāv*. The Tamil transliteration given is **நி:நாவ்** *nHīnāv*, apparently with the wrong vowel in the first syllable, and with the colon-like mark. The Tamil translation given is **வழுக்கை** *vazakkai* ‘bald’. The Hindi word for this is given (in Tamil script!) as **க<sup>3</sup>ஞ்ஜா** *ka<sup>3</sup>ñjā*, i.e., *gañjā* ‘bald’. In L2/03-256 Ka’ōnohi Kai has suggested that the aspirates be encoded as digraphs, but this seems unnecessary. In some implementations, considerable space can be seen between the liquid or nasal and the UPAKSHARA, but this is probably an artifact of font technology.

In a document in 2003 (L2/03-277), Peri Bhaskararao suggested that the UPAKSHARA could be analyzed as a combining form of HA. No additional information in support of that position has been provided to overturn the analysis we have given above. Though the chart in *Bhāṣābhīmānī* (shown in Figure 7) gives ற் rh as the killed Tamil equivalent to Saurashtra ౠ rh, transliterations elsewhere would suggest ற: rh. For Saurashtra to be equivalent to ற் rh, we might expect to see ౠ rh, that is, killed ra + killed ha. The equivalent Devanagari sequence would be र्ह rh with two viramas. We believe that the evidence suggests that ౠ rh (with killed vowel) and ౠ rha (without killed vowel, equivalent to र्ह rha) would be distinct in the Saurashtra script. Consider the following transliteration opportunity. A Sanskrit phrase ओम् सिमंहा कणपतये नमः Om simha Kaṇapataye namaha appears in Tamil script on a web page at <http://srripambattisiddhar.com/tamil/poojai/vinayaga.htm>; note the consonant cluster ம்ஹ mha. In Saurashtra this could be

written ୠ ୠୠୢ ୠୠୢୢୢ ୠୢ: ୠୢ: *Om simha Gaṇapataye namaḥ*, assuming that Sanskrit *mha* is not the aspirate *mHa*. Contrast this with the pair ୠୢ:ଁ ୠୢ:ଁ *mu:ri* and ୠୢୢୢୢ *mHuri*—both examples from *Bhāṣābhīmānī*. (We are aware that this argument would be stronger if we had an actual *pūja* text in Saurashtran script to show the transliteration practice for Sanskrit.) In any case, now, in 2005, we reiterate our proposal for UPAKSHARA to be encoded as a unique Saurashtran character.

### Conjunct characters

As noted above, the virama is by default visible in Saurashtra, and ZWJ is added to force a conjunct; ZWNJ will also force a visible virama:

ଁ KA	+	ୠ	+			ୡ SSA	=	ଁୡ <i>kṣa</i>
ଁ KA	+	ୠ	+	ZWJ	+	ୡ SSA	=	ଁୡ <i>kṣa</i>
ଁ KA	+	ୠ	+	ZWNJ	+	ୡ SSA	=	ଁୡ <i>kṣa</i>

Older Saurashtra texts have not been available to us, but the following conjunct forms are known as attested in Figure 2 below:

ୢ SHA	+	ୠ	+			ୡ RA	=	ୢୡ <i>śra</i>						
ୢ SHA	+	ୠ	+	ZWJ	+	ୡ RA	=	ୢୡ <i>śra</i>						
ୡ SSA	+	ୠ	+			୦ TTHA	=	ୡ୦ <i>ṣṭha</i>						
ୡ SSA	+	ୠ	+	ZWJ	+	୦ TTHA	=	ୡ୦ <i>ṣṭha</i>						
ଁ TA	+	ୠ	+			ୡ YA	=	ଁୡ <i>tya</i>						
ଁ TA	+	ୠ	+	ZWJ	+	ୡ YA	=	ଁୡ <i>tya</i>						
ୡ SSA	+	ୠ	+			ୢ TTA	+	ୠ	+			ୡ RA	=	ୡୢୡ <i>ṣṭra</i>
ୡ SSA	+	ୠ	+	ZWJ	+	ୢ TTA	+	ୠ	+	ZWJ	+	ୡ RA	=	ୡୢୡ <i>ṣṭra</i>

An implementation of older Saurashtra will entail further study as to the specific consonant clusters used.

### Implementations

Ka'ōnohi Kai implemented Saurashtra for Mac OS X 10.2 in 2003, based on a font by Jason Glavy. This commercial product was based on my exploratory Unicode Saurashtra proposal (published at [www.evertype.com/standards/iso10646/pdf/saurashtra.pdf](http://www.evertype.com/standards/iso10646/pdf/saurashtra.pdf)). The encoding model and repertoire used by XenoType is a subset of the model used in *this* proposal; it is clear that the model works well for Saurashtra. There are also two TrueType implementations of the modern Saurashtra script in circulation on the internet. To date, no computer implementation of the older version of the script has been created. Previously, a letterpress was used to print *Bhāṣābhīmānī* magazine. As printing technology advanced, the need for a Saurashtra script font was felt, and was created by local publishers. The *Suresh* font, designed by Suresh Thimma Ramdas, is used to print *Bhāṣābhīmānī*; a later corrected version of that font is *Kuber*, which is used here. (The name *Kuber* is derived from “Kubendiran”, the name of the editor of this magazine.) Traditional implementation in lead type was known from the end of the 19th century. According to Norihiko Učida:

Saurashtrans had their own beautiful and systematic script, the origin of which is not known. At the end of the last century T. M. Rama Ray reformed the script and cast types for it. He published many books in

this script. Later he introduced some changes and standardized it. which is now used by some people as Saurashtra script. Since the end of the nineteenth century books have been printed in Telugu, Tamil, Devanagari and Saurashtra scripts. When Gandhiji preached for one common script for India, one group of the community decided to abandon the Saurashtra script in favour of Devanagari and destroyed the matrix designed by Rama Ray.

## Collating order

The usual Sanskrit ordering applies to Saurashtra: A AA I II U UU RU RUU LU LUU E EE AI O OO AU AM AH UPAKSHARA KA KHA GA GHA NGA CA CHA JA JHA NYA TTA TTHA DDA DDHA NNA TA THA DA DHA NA PA PHA BA BHA MA YA RA LA VA SHA SSA SA HA LLA KSSA

## Unicode Character Properties

The Saurashtra characters have properties similar to Devanagari and the other Indic scripts. Since Saurashtra uses no truly non-spacing marks, all of the combining classes are zero, with the exception of the virama. All of the letters are in the Letter other category, and have strong left-to-right directionality. The vowel signs all follow the consonants with which they are associated, and therefore are combining class zero as well, analogous to the Devanagari model. The virama has the same combining class (9) as other virama characters. The digits likewise are analogous to the Devanagari and other digits, with the same properties, including decimal digit values. The table below shows all of the properties defined in the UnicodeData.txt file.

```
A880;SAURASHTRA SIGN ANUSVARA;Mc;0;L;;;;N;;;;;
A881;SAURASHTRA SIGN VISARGA;Mc;0;L;;;;N;;;;;
A882;SAURASHTRA LETTER A;Lo;0;L;;;;N;;;;;
A883;SAURASHTRA LETTER AA;Lo;0;L;;;;N;;;;;
A884;SAURASHTRA LETTER I;Lo;0;L;;;;N;;;;;
A885;SAURASHTRA LETTER II;Lo;0;L;;;;N;;;;;
A886;SAURASHTRA LETTER U;Lo;0;L;;;;N;;;;;
A887;SAURASHTRA LETTER UU;Lo;0;L;;;;N;;;;;
A888;SAURASHTRA LETTER VOCALIC R;Lo;0;L;;;;N;;;;;
A889;SAURASHTRA LETTER VOCALIC RR;Lo;0;L;;;;N;;;;;
A88A;SAURASHTRA LETTER VOCALIC L;Lo;0;L;;;;N;;;;;
A88B;SAURASHTRA LETTER VOCALIC LL;Lo;0;L;;;;N;;;;;
A88C;SAURASHTRA LETTER E;Lo;0;L;;;;N;;;;;
A88D;SAURASHTRA LETTER EE;Lo;0;L;;;;N;;;;;
A88E;SAURASHTRA LETTER AI;Lo;0;L;;;;N;;;;;
A88F;SAURASHTRA LETTER O;Lo;0;L;;;;N;;;;;
A890;SAURASHTRA LETTER OO;Lo;0;L;;;;N;;;;;
A891;SAURASHTRA LETTER AU;Lo;0;L;;;;N;;;;;
A892;SAURASHTRA LETTER KA;Lo;0;L;;;;N;;;;;
A893;SAURASHTRA LETTER KHA;Lo;0;L;;;;N;;;;;
A894;SAURASHTRA LETTER GA;Lo;0;L;;;;N;;;;;
A895;SAURASHTRA LETTER GHA;Lo;0;L;;;;N;;;;;
A896;SAURASHTRA LETTER NGA;Lo;0;L;;;;N;;;;;
A897;SAURASHTRA LETTER CA;Lo;0;L;;;;N;;;;;
A898;SAURASHTRA LETTER CHA;Lo;0;L;;;;N;;;;;
A899;SAURASHTRA LETTER JA;Lo;0;L;;;;N;;;;;
A89A;SAURASHTRA LETTER JHA;Lo;0;L;;;;N;;;;;
A89B;SAURASHTRA LETTER NYA;Lo;0;L;;;;N;;;;;
A89C;SAURASHTRA LETTER TTA;Lo;0;L;;;;N;;;;;
A89D;SAURASHTRA LETTER TTHA;Lo;0;L;;;;N;;;;;
A89E;SAURASHTRA LETTER DDA;Lo;0;L;;;;N;;;;;
A89F;SAURASHTRA LETTER DDHA;Lo;0;L;;;;N;;;;;
A8A0;SAURASHTRA LETTER NNA;Lo;0;L;;;;N;;;;;
A8A1;SAURASHTRA LETTER TA;Lo;0;L;;;;N;;;;;
A8A2;SAURASHTRA LETTER THA;Lo;0;L;;;;N;;;;;
A8A3;SAURASHTRA LETTER DA;Lo;0;L;;;;N;;;;;
A8A4;SAURASHTRA LETTER DHA;Lo;0;L;;;;N;;;;;
A8A5;SAURASHTRA LETTER NA;Lo;0;L;;;;N;;;;;
A8A6;SAURASHTRA LETTER PA;Lo;0;L;;;;N;;;;;
A8A7;SAURASHTRA LETTER PHA;Lo;0;L;;;;N;;;;;
A8A8;SAURASHTRA LETTER BA;Lo;0;L;;;;N;;;;;
```

A8A9;SAURASHTRA LETTER BHA;Lo;0;L;;;;N;;;;  
A8AA;SAURASHTRA LETTER MA;Lo;0;L;;;;N;;;;  
A8AB;SAURASHTRA LETTER YA;Lo;0;L;;;;N;;;;  
A8AC;SAURASHTRA LETTER RA;Lo;0;L;;;;N;;;;  
A8AD;SAURASHTRA LETTER LA;Lo;0;L;;;;N;;;;  
A8AE;SAURASHTRA LETTER VA;Lo;0;L;;;;N;;;;  
A8AF;SAURASHTRA LETTER SHA;Lo;0;L;;;;N;;;;  
A8B0;SAURASHTRA LETTER SSA;Lo;0;L;;;;N;;;;  
A8B1;SAURASHTRA LETTER SA;Lo;0;L;;;;N;;;;  
A8B2;SAURASHTRA LETTER HA;Lo;0;L;;;;N;;;;  
A8B3;SAURASHTRA LETTER LLA;Lo;0;L;;;;N;;;;  
A8B4;SAURASHTRA CONSONANT SIGN UPAKSHARA;Mc;0;L;;;;N;;;;  
A8B5;SAURASHTRA VOWEL SIGN AA;Mc;0;L;;;;N;;;;  
A8B6;SAURASHTRA VOWEL SIGN I;Mc;0;L;;;;N;;;;  
A8B7;SAURASHTRA VOWEL SIGN II;Mc;0;L;;;;N;;;;  
A8B8;SAURASHTRA VOWEL SIGN U;Mc;0;L;;;;N;;;;  
A8B9;SAURASHTRA VOWEL SIGN UU;Mc;0;L;;;;N;;;;  
A8BA;SAURASHTRA VOWEL SIGN VOCALIC R;Mc;0;L;;;;N;;;;  
A8BB;SAURASHTRA VOWEL SIGN VOCALIC RR;Mc;0;L;;;;N;;;;  
A8BC;SAURASHTRA VOWEL SIGN VOCALIC L;Mc;0;L;;;;N;;;;  
A8BD;SAURASHTRA VOWEL SIGN VOCALIC LL;Mc;0;L;;;;N;;;;  
A8BE;SAURASHTRA VOWEL SIGN E;Mc;0;L;;;;N;;;;  
A8BF;SAURASHTRA VOWEL SIGN EE;Mc;0;L;;;;N;;;;  
A8C0;SAURASHTRA VOWEL SIGN AI;Mc;0;L;;;;N;;;;  
A8C1;SAURASHTRA VOWEL SIGN O;Mc;0;L;;;;N;;;;  
A8C2;SAURASHTRA VOWEL SIGN OO;Mc;0;L;;;;N;;;;  
A8C3;SAURASHTRA VOWEL SIGN AU;Mc;0;L;;;;N;;;;  
A8C4;SAURASHTRA SIGN VIRAMA;Mn;9;NSM;;;;N;;;;  
A8CE;SAURASHTRA DANDA;Po;0;L;;;;N;;;;  
A8CF;SAURASHTRA DOUBLE DANDA;Po;0;L;;;;N;;;;  
A8D0;SAURASHTRA DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;  
A8D1;SAURASHTRA DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;  
A8D2;SAURASHTRA DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;  
A8D3;SAURASHTRA DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;  
A8D4;SAURASHTRA DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;  
A8D5;SAURASHTRA DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;  
A8D6;SAURASHTRA DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;  
A8D7;SAURASHTRA DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;  
A8D8;SAURASHTRA DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;  
A8D9;SAURASHTRA DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;

## References

Norihiko Učida. *Language of the Saurashtrans in Tirupati*. (In Latin script.)

Norihiko Učida. *Saurashtra-English dictionary*. ISBN 3447030550. (In Latin script.) (Učida is a Japanese scholar who went to Madurai to research Saurashtra. He did his research at Madurai Kamarajar University. The local language in Madurai is Tamil; in the course of his research, he learned Saurashtra well and is said to speak Saurashtra as well as a native Saurashtrian.)

பதார்த்தன் *Padārthun*. 1998. A magazine published in Madurai.

பிஹாஷ் பிம்னி *Bhāṣābhīmāni*. 2001. A magazine published in Madurai.

சுப்ரமணியன், ஸ்ரீ தாடா. 2000. *ஸவ்ராஷ்ட்ர பாத்<sup>2</sup> மாலொ*. மதுரை: ப்ரியா பதப்பகம் ப்ரியா டெக்னிகல் இன்ஸ்டிடியூட். [Śrī Tāṭā. Cupramaṇiyaṇ. 2000. *Savrāṣṭra pātha mālo* = *Savrāṣṭra pāṭ<sup>2</sup>a mālo*. Madurai: Priyā Patappakam Priyā Ṭekṇikal Inṣṭīyut]

## Acknowledgements

This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley).

## Figures

உயிரெழுத்துக்கள்

அ	ஆ	இ	ஈ	உ	ஊ	஋	஌	஍
அ	ஆ	இ	ஈ	உ	ஊ	஋	஌	஍
ஏ	ஐ	ஒ	ஔ	஥	஦	஧	ந	ன
ஏ	ஐ	ஒ	ஔ	஥	஦	஧	ந	ன

**Figure 1.** Saurashtra vowels, with equivalents in Tamil and Telugu scripts.

[illegible]

**Figure 2.** The text in the bottom half of this figure is given in Saurashtra and Tamil script. Note the unusual forms of  $-l$  and  $-\bar{l}$ . The Saurashtra in the second line is in an early form of the script as it shows conjunct consonants. It is rendered here with modern orthography and transliteration:

මුදල, පෞරුෂය, පැරණි පැරණි, පැරණි, පැරණි |

ඉන්ද්‍රියං භක්තයිත්‍රං භජේත් භග්, භුක්තෑ |

śrēṣṭha saurāṣṭra sāhitya sabhā, machuro |

ஸ்ரேஷ்ட ஸௌராஷ்ட்ர ஸாஹீத்ய ஸபா, மதுரை

srēṣṭa saurāṣṭra sāhītya sapā, maturai

**அஜாது**

[illegible]

**ஒடிது**

॥ <sup>क<sup>1</sup></sup> क	ॡ <sup>क<sup>2</sup></sup> ख	ॢ <sup>क<sup>3</sup></sup> ग	ॣ <sup>क<sup>4</sup></sup> घ	। <sup>ख</sup> ङ
॥ <sup>च<sup>1</sup></sup> च	ॡ <sup>च<sup>2</sup></sup> छ	ॢ <sup>ज</sup> ज	ॣ <sup>झ</sup> झ	। <sup>ञ</sup> ञ
॥ <sup>ट<sup>1</sup></sup> ट	ॡ <sup>ठ</sup> ठ	ॢ <sup>ड</sup> ड	ॣ <sup>ढ</sup> ढ	। <sup>ण</sup> ण
॥ <sup>त<sup>1</sup></sup> त	ॡ <sup>थ</sup> थ	ॢ <sup>द</sup> द	ॣ <sup>ध</sup> ध	। <sup>न</sup> न
॥ <sup>प<sup>1</sup></sup> प	ॡ <sup>फ</sup> फ	ॢ <sup>ब</sup> ब	ॣ <sup>भ</sup> भ	। <sup>म</sup> म
॥ <sup>य</sup> य	ॡ <sup>र</sup> र	ॢ <sup>ल</sup> ल	ॣ <sup>व</sup> व	। <sup>श</sup> श
॥ <sup>ष</sup> ष	ॡ <sup>स</sup> स	ॢ <sup>ह</sup> ह	ॣ <sup>ळ</sup> ळ	। <sup>क्ष</sup> क्ष

[illegible]


Page 7



தநுநு	
க ன க	க்லா னநு -
கா ன7 கா	கெ ன9 -
கி ன் கி	கே ன8 கெ
கீ னி கி	கை ன9 கெ
கு ன8 கு	கொ ன2 -
கூ ன8 கூ	கோ ன4 கொ
க்ரு ன5 க்	கௌ ன3 கௌ
க்ரு ன5 -	கம் ன0 க்
க்லு ன0 -	கஹ ன1 க:

**Figure 5.** Sample from the magazine *Padārthun*, showing Saurashtra vowel matras with equivalents in Tamil and Devanagari scripts.

**ஸவ்ராஷ்ட்ர மாதா  
ஸஃஸ7புஷ்ட ஸ707**

1. ஸெ ஸெ ஸஃஸ7புஷ்ட ஸ707   
ஸெ ஸெ ஸஃஸ7புஷ்ட ஸ707
  2. ஸஃஸ7 உஃஸ7ஸ7 ஸ7 ஸ7 ஸ7 ஸ7  
ஸ7ஸ7 ஸ7ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7  
ஸ7ஸ7 ஸ7ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7  
ஸ7ஸ7 ஸ7ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 (ஸெஸெ) ||
  3. ஸ7ஸ7ஸ7 ஸ7ஸ7 ஸ7ஸ7 ஸ7ஸ7 ஸ7ஸ7  
ஸ7ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7  
ஸ7ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7  
ஸ7ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 ஸ7 (ஸெஸெ) ||
- 
- 1 ஜய ஜய ஸவ்ராஷ்ட்ர மாதா,  
ஜய ஜய ஸவ்ராஷ்ட்ர மாதா.
  - 2 ஸோஷ தேஷுஷுக் தூ ராணி மாதா  
ஸுமிது குலம் தியெ ஸவ்ராஷ்ட்ர மாதா.  
ஸோஷ ஜெஹும் க2டி3 தொகொ நமு மாதா  
ஸொந்தம்மு ப2ய்லொ தூஸ் ஸவ்ராஷ்ட்ர மாதா (ஜய ஜய)
  - 3 காவ்யமு பூ4ஷுஷுன் ஸஸர் ஸஸர் கெரிகின்  
க4லுஸு மீ தொகொ ஸவ்ராஷ்ட்ர மாதா.  
ஜீவ்யமு லெக்கு யே, தெ4ய்ரியம் ஜு2க்கு ஸே  
ஜெக3துரு ப2ய்லொ தூஸ் ஸவ்ராஷ்ட்ர மாதா (ஜய ஜய)

**Figure 6.** Sample from the magazine *Bhāṣābhīmāni*, showing a Saurashtra poem with both SAURASHTRA DANDA and SAURASHTRA DOUBLE DANDA. A Tamil-script transliteration of the poem is also given.

## ரோமன் எழுத்து - ஸௌராஷ்டிர மொழி

- ஜாபாலி

ரோமன் (ஆங்கில) எழுத்துகளைக் கொண்டு ஸௌராஷ்டிர மொழியை எழுதப் பயன்படுத்தும் முறை வெளிநாடுகளில் வாழும் ஸௌராஷ்டிரர்கள் இம் முறையினைப் பயன்படுத்தி தம் உறவினர்களுக்கும் நண்பர்களுக்கும் கடிதம் மற்றும் E-mailகளை அனுப்பலாம். ஏன், இங்குள்ளவர்களும் வயது முதிர்ந்த - புது வகை வரி வடிவத்தினைக் கற்றுக் கொள்ள முடியாது என்று நம்புவர்கள் - இதனைப் பயன்படுத்தலாமே.

### VOWELS - உயிரெழுத்துகள்

ஐ ஐ ஐ ஐ ஐ ஐ ஐ ஐ ஐ ஐ  
அ ஆ இ ஈ உ ஊ எ ஏ ஒ ஓ  
A Ā I Ī U Ū E Ē O Ō

### CONSONANTS - மெய்யெழுத்துகள்

க	க்	க்	க்	ங்	ச	ச்	ஜ்	ஜ்	ஞ்
K	Kh	G	Gh	Ng	C	Ch	J	Jh	Ņg
த	த்	த்	த்	ண்	த	த்	த்	த்	ந்
T	Th	D	Dh	Ṇ	T	Th	D	Dh	N
ப	ப்	ப்	ப்	ம்	ய்	ர	ல்	வ்	ஸ்
P	Ph	B	Bh	M	Y	R	L	V	Ś
ஷ	ஸ்	ஹ்	ள்	க்ஷ	ர்ஹ்	ல்ஹ்	ம்ஹ்	ந்ஹ்	
Sh	S	H	L	Ksh	Rh	Lh	Mh	Nh	

### CONSONANT ADDED VOWELS - உயிர் மெய்யெழுத்துகள்

க	கா	கி	கீ	கு	கூ	கெ	கே	கொ	கோ
KA	KĀ	KI	KĪ	KU	KŪ	KE	KĒ	KO	KŌ

### EXAMPLES - உதாரணங்கள்

ப்ஷாபிமானி	இயெ அநுசரி அம்தெநு லேகூந் லிக்குவொ
Bhāśābhīmāni	IYE ANUCARI AMTENU LĒKHUN LIKKUVO
ஸவ்ராஷ்டிர மணி	ஒண்டெதிம் ஸிக்குல்டுவாயி
SAVRĀSHĪTRA MAṆI	ONTEḌIM SIKKULḌUVĀYI

தீபாவளி மலர் 2001

Figure 7. Sample from the magazine *Bhāśābhīmāni*, showing Saurashtra letters with equivalents in Tamil and Latin. The consonants are shown with their vowels killed; the UPAKSHARA character is shown in combination with killed consonants, indicated with the arrow. Below that the vowel matras are given in combination with consonant ௧ ka.



Figure 8. Sample from the magazine *Bhāṣābhīmānī*, showing Saurashtra prose text. Note the use of SAURASHTRA DANDA.

Row A8: SAURASHTRA

	A88	A89	A8A	A8B	A8C	A8D
0	◌◌	ଐ	ଐ	ଐ	ଐ	୦
1	◌ଃ	ଐ	ଠ	ଠ	ଠ	୧
2	ଐ	ଠ	ଐ	ଐ	ଐ	୨
3	ଐ	ଠ	ଐ	ଐ	ଐ	୩
4	ଐ	ଠ	ଐ	ଐ	ଐ	୪
5	ଐ	ଐ	ଐ	ଐ		୫
6	ଐ	ଐ	ଐ	ଐ		୬
7	ଐ	ଐ	ଐ	ଐ		୭
8	ଐ	ଐ	ଐ	ଐ		୮
9	ଐ	ଐ	ଐ	ଐ		୯
A	ଐ	ଐ	ଐ	ଐ		
B	ଐ	ଐ	ଐ	ଐ		
C	ଐ	ଐ	ଐ	ଐ		
D	ଐ	ଐ	ଐ	ଐ		
E	ଐ	ଐ	ଐ	ଐ		
F	ଐ	ଐ	ଐ	ଐ		

G = 00  
P = 00

## Row A8: SAURASHTRA

hex	Name	hex	Name
A880	SAURASHTRA SIGN ANUSVARA	A8D9	SAURASHTRA DIGIT NINE
A881	SAURASHTRA SIGN VISARGA	A8DA	(This position shall not be used)
A882	SAURASHTRA LETTER A	A8DB	(This position shall not be used)
A883	SAURASHTRA LETTER AA	A8DC	(This position shall not be used)
A884	SAURASHTRA LETTER I	A8DD	(This position shall not be used)
A885	SAURASHTRA LETTER II	A8DE	(This position shall not be used)
A886	SAURASHTRA LETTER U	A8DF	(This position shall not be used)
A887	SAURASHTRA LETTER UU		
A888	SAURASHTRA LETTER VOCALIC R		
A889	SAURASHTRA LETTER VOCALIC RR		
A88A	SAURASHTRA LETTER VOCALIC L		
A0BB	SAURASHTRA LETTER VOCALIC LL		
A88C	SAURASHTRA LETTER E		
A88D	SAURASHTRA LETTER EE		
A88E	SAURASHTRA LETTER AI		
A88F	SAURASHTRA LETTER O		
A890	SAURASHTRA LETTER OO		
A891	SAURASHTRA LETTER AU		
A892	SAURASHTRA LETTER KA		
A893	SAURASHTRA LETTER KHA		
A894	SAURASHTRA LETTER GA		
A895	SAURASHTRA LETTER GHA		
A896	SAURASHTRA LETTER NGA		
A897	SAURASHTRA LETTER CA		
A898	SAURASHTRA LETTER CHA		
A899	SAURASHTRA LETTER JA		
A89A	SAURASHTRA LETTER JHA		
A89B	SAURASHTRA LETTER NYA		
A89C	SAURASHTRA LETTER TTA		
A89D	SAURASHTRA LETTER TTHA		
A89E	SAURASHTRA LETTER DDA		
A89F	SAURASHTRA LETTER DDHA		
A8A0	SAURASHTRA LETTER NNA		
A8A1	SAURASHTRA LETTER TA		
A8A2	SAURASHTRA LETTER THA		
A8A3	SAURASHTRA LETTER DA		
A8A4	SAURASHTRA LETTER DHA		
A8A5	SAURASHTRA LETTER NA		
A8A6	SAURASHTRA LETTER PA		
A8A7	SAURASHTRA LETTER PHA		
A8A8	SAURASHTRA LETTER BA		
A8A9	SAURASHTRA LETTER BHA		
A8AA	SAURASHTRA LETTER MA		
A8AB	SAURASHTRA LETTER YA		
A8AC	SAURASHTRA LETTER RA		
A8AD	SAURASHTRA LETTER LA		
A8AE	SAURASHTRA LETTER VA		
A8AF	SAURASHTRA LETTER SHA		
A8B0	SAURASHTRA LETTER SSA		
A8B1	SAURASHTRA LETTER SA		
A8B2	SAURASHTRA LETTER HA		
A8B3	SAURASHTRA LETTER LLA		
A8B4	SAURASHTRA CONSONANT SIGN UPAKSHARA		
A8B5	SAURASHTRA VOWEL SIGN AA		
A8B6	SAURASHTRA VOWEL SIGN I		
A8B7	SAURASHTRA VOWEL SIGN II		
A8B8	SAURASHTRA VOWEL SIGN U		
A8B9	SAURASHTRA VOWEL SIGN UU		
A8BA	SAURASHTRA VOWEL SIGN VOCALIC R		
A8BB	SAURASHTRA VOWEL SIGN VOCALIC RR		
A8BC	SAURASHTRA VOWEL SIGN VOCALIC L		
A8BD	SAURASHTRA VOWEL SIGN VOCALIC LL		
A8BE	SAURASHTRA VOWEL SIGN E		
A8BF	SAURASHTRA VOWEL SIGN EE		
A8C0	SAURASHTRA VOWEL SIGN AI		
A8C1	SAURASHTRA VOWEL SIGN O		
A8C2	SAURASHTRA VOWEL SIGN OO		
A8C3	SAURASHTRA VOWEL SIGN AU		
A8C4	SAURASHTRA SIGN VIRAMA		
A8C5	(This position shall not be used)		
A8C6	(This position shall not be used)		
A8C7	(This position shall not be used)		
A8C8	(This position shall not be used)		
A8C9	(This position shall not be used)		
A8CA	(This position shall not be used)		
A8CB	(This position shall not be used)		
A8CC	(This position shall not be used)		
A8CD	(This position shall not be used)		
A8CE	SAURASHTRA DANDA		
A8CF	SAURASHTRA DOUBLE DANDA		
A8D0	SAURASHTRA DIGIT ZERO		
A8D1	SAURASHTRA DIGIT ONE		
A8D2	SAURASHTRA DIGIT TWO		
A8D3	SAURASHTRA DIGIT THREE		
A8D4	SAURASHTRA DIGIT FOUR		
A8D5	SAURASHTRA DIGIT FIVE		
A8D6	SAURASHTRA DIGIT SIX		
A8D7	SAURASHTRA DIGIT SEVEN		
A8D8	SAURASHTRA DIGIT EIGHT		

## A. Administrative

### 1. Title

Final revised proposal to encode the Saurashtra script in the UCS.

### 2. Requester's name

Michael Everson and Jeyakumar Chinnakkonda Krishnamoorthy

### 3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

### 4. Submission date

2005-09-21

### 5. Requester's reference (if applicable)

### 6. Choose one of the following:

#### 6a. This is a complete proposal

Yes.

#### 6b. More information will be provided later

No.

## B. Technical – General

### 1. Choose one of the following:

#### 1a. This proposal is for a new script (set of characters)

Yes.

#### Proposed name of script

Saurashtra.

#### 1b. The proposal is for addition of character(s) to an existing block

No.

#### 1b. Name of the existing block

#### 2. Number of characters in proposal

79

#### 3. Proposed category (see section II, Character Categories)

Category A.

#### 4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)

Level 2.

#### 4b. Is a rationale provided for the choice?

Yes.

#### 4c. If YES, reference

Brahmic script with vowel signs.

#### 5a. Is a repertoire including character names provided?

Yes.

#### 5b. If YES, are the names in accordance with the naming guidelines in Annex L of ISO/IEC 10646-1: 2000?

Yes.

#### 5c. Are the character shapes attached in a legible form suitable for review?

Yes.

#### 6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Jeyakumar C. K. (www.palkar.org) & Michael Everson (Evince). TrueType.

#### 6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Jeyakumar C. K. (www.palkar.org) & Michael Everson (Evince). Fontographer.

#### 7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

#### 7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

#### 8. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

#### 9. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.

See Unicode properties above.

## C. Technical – Justification

### 1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Yes. See N2607.

### 2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

### 2b. If YES, with whom?

Jeyakumar C. K. is a user himself, and has been in touch with Mr Kubendiran, editor of ્ Bhāṣābhīmānī.

**2c. If YES, available relevant documents**

N/A.

**3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?**

Yes.

**4a. The context of use for the proposed characters (type of use; common or rare)**

Saurashtra script is used fairly rarely, but it is enjoying a renaissance, and certainly has some commercial value (there are even advertisements in the monthly magazine *Bhāṣābhīmāni* for “Desk Top Publishing in English, Tamil, Hindi, Malayalam, Saurashtra”). Both the Saurashtran and Tamil scripts are used to write Saurashtra. Use of Saurashtra script is primarily cultural and religious, though some private schools are now teaching the script to children.

**4b. Reference**

**5a. Are the proposed characters in current use by the user community?**

Yes.

**5b. If YES, where?**

In the Indian states Tamil Nadu and Andhra Pradesh. The monthly magazine *Bhāṣābhīmāni* is published in Madurai. Much of this magazine is written in Tamil and in Saurashtra in Tamil script, but Saurashtra script is also used in it.

**6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?**

Yes.

**6b. If YES, is a rationale provided?**

Yes.

**6c. If YES, reference**

Accordance with the Roadmap.

**7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?**

Yes.

**8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?**

No. Superficial

**8b. If YES, is a rationale for its inclusion provided?**

**8c. If YES, reference**

**9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?**

No.

**9b. If YES, is a rationale for its inclusion provided?**

**9c. If YES, reference**

**10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?**

No.

**10b. If YES, is a rationale for its inclusion provided?**

**10c. If YES, reference**

**11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?**

Yes.

**11b. If YES, is a rationale for such use provided?**

Yes.

**11c. If YES, reference**

Brahmic vowel signs.

**12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?**

No.

**12b. If YES, reference**

**13a. Does the proposal contain characters with any special properties such as control function or similar semantics?**

No.

**13b. If YES, describe in detail (include attachment if necessary)**

**14a. Does the proposal contain any Ideographic compatibility character(s)?**

No.

**14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?**

**14c. If YES, reference**